Designing Novel Multi-View Clustering Algorithms for Enhanced Data Integration and Clustering Performance

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

to

Manav Rachna International Institute of Research and Studies (Deemed to be University)

by

Pankaj Kumar
23-24/Ph.D/S/025

Under the supervision of

Dr. Rashmi Agrawal
Professor, School of Computer Application

School of Computer Application
Manav Rachna International Institute of Research & Studies Sector-43, Surajkund – Delhi Road, Faridabad – 121004

SCHOOL OF COMPUTER APPLICATION

Certificate

I hereby declare that the Short Synopsis titled "**Designing Novel Multi-View Clustering Algorithms for Enhanced Data Integration and Clustering Performance**" being submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the Department of Computer Applications under School of Computer Applications of Manav Rachna International Institute of Research & Studies, Faridabad during the academic year 2024 – 2025. It is a bona fide record of original work carried out under the supervision of **Dr. Rashmi Agrawal** and has not been submitted in full or part to any other university or institute for a degree or diploma. Further, it is certified that Short Synopsis is plagiarism free, and proof has been submitted.

Pankaj Kumar
23-24/Ph.D/S/025

Signature of the Supervisor
Dr. Rashmi Agrawal, Professor
School of Computer Applications (MRIIRS)

(Chairperson, DRC)

Dean,
School of Computer Applications
(MRIIRS)

ABSTRACT

Clustering is a fundamental task in data analysis and machine learning, where the goal is to group similar data points based on their inherent characteristics. Traditional clustering methods, such as k-means and hierarchical clustering, operate on single-view data, assuming all the information required to cluster the data is available in one form. However, in many real-world applications, data often comes from multiple sources or perspectives, referred to as views, which present unique challenges. One major issue with traditional clustering is its inability to capture the full complexity of multi-represented data, leading to suboptimal clustering results.

Multi-view clustering (MVC) addresses this issue by integrating data from multiple views and leveraging complementary information to produce more accurate and meaningful clusters. MVC has become essential in bioinformatics, image processing, and social network analysis, where data is naturally multi-faceted. The need for MVC arises from the limitations of single-view clustering, such as its lack of robustness when handling incomplete or noisy data, and its failure to fully utilize the richness of data from different views. MVC techniques aim to resolve these challenges by integrating diverse information while maintaining computational efficiency and scalability. Some prominent MVC methods include Canonical Correlation Analysis (CCA), Spectral Clustering, Non-negative Matrix Factorization (NMF), and Co-regularized Clustering.

Despite its advantages, current MVC approaches face significant challenges, including the effective handling of missing data, robustly integrating views, and scaling algorithms for large datasets. This research aims to address these gaps by critically reviewing existing MVC techniques, developing novel algorithms to handle missing data, and designing enhanced integration methods using distributed computing, regularization, and probabilistic modeling. Furthermore, the performance of the proposed algorithms will be rigorously evaluated based on key metrics like clustering accuracy, computational efficiency, scalability, and robustness. By addressing these objectives, this research seeks to advance the state-of-the-art in MVC, providing more effective solutions for real-world multi-view

data clustering problems.

Keywords: Clustering, Deep Learning, Multi-view Clustering, Multi-view K-means, Multi-view Matrix Factorization, Multi-view Spectral Clustering

Table of Contents

Certificate		i
Abstract		ii
Table of Contents		iv

S.No.	Description	Page No.
1	Introduction	1
	1.1 The Emergence of Multi-View Clustering	2
	1.2 Multi-View Clustering Methods	2
2	Literature Review and Research Gaps	3
	2.1 Research Gaps	14
3	Problem Statement	15
4	Description of Broad Area	16
5	Research Objective	18
6	Research Methodology	19
7	Expected Outcome	20
8	References	20

1. INTRODUCTION

Nowadays it is believed that the availability of various comprehensive sets of data enlarges the range of data analysis opportunities. Accordingly, clustering pays much attention—a prominent approach to unsupervised learning—among them. Cluster analysis seeks to put closely related data points in the same group with the hope of uncovering the natural data patterns in it. Most conventional clustering techniques work with one vision of the data which in most cases captures only a fraction of all the data's richness. However, many real-world datasets are naturally multi-view (Figure 1), it is a case where besides a set of features we have another set or perspective that describes the same objects.

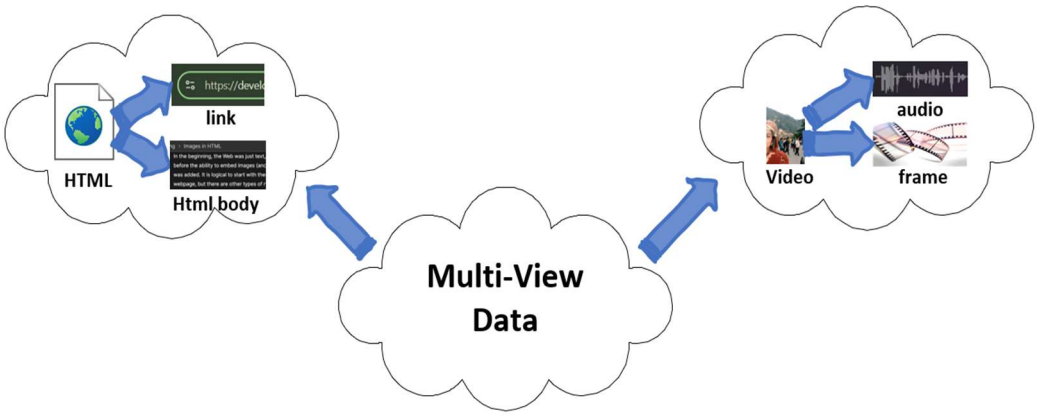

Figure 1: Multi-View Data (Source- Self)

Clustering has been used in data analysis for many years, some of the traditional clustering types include k-means, hierarchical, and Gaussian mixture models. These approaches have been well-researched and used in numerous fields including bioinformatics, and social network analysis. However, the existing clustering methods assume that the point that the data is subsided in a single feature space while multi-view data provides more opportunities to be utilized. With an increase in the availability and heterogeneity of the data sources, it is increasingly apparent that the single-view clustering technique is insufficient.

Some of the challenges of multi-view data have led to the introduction of multi-view learning where both supervised and unsupervised learning tasks belong. Multi-View learning uses multiple sets of features known as views to enhance the performance

of learning. In the context of clustering Multi-View Learning led to the development of the Multiple-View Clustering technique; the idea of which is to cluster the data for the information given in all views.

1.1 THE EMERGENCE OF MULTI-VIEW CLUSTERING

More formal study, of Multi-View Clustering can be traced back to 2004 when Bickel and Scheffer [1] proposed the Multi-View K-means and Multi-View expectation maximization (EM) methods. Their work showed that using several views could amount to a huge improvement in clustering performance. Especially, when the several views contain different information about the inherent structure of the data. Since then, several extended versions of MVC algorithms have been developed to cope with the difficulties that arise with Multi-View data.

1.2 MULTI-VIEW CLUSTERING METHODS

Bickel and Scheffer [1] were pioneered in studying Multi-View clustering during 2004, in his research they adapted the classic K-Means & expectation maximization (EM) clustering methods to experiment with text data which were two conditionally independent view. He laid the foundation on which many other methods were developed over the past 20 years. Among all methods there are five popular clustering methods that has been extended/adapted for Multi-View methods (Figure 2): k-means, spectral clustering, matrix factorization, tensor decomposition, and deep learning. These five algorithms can be used to integrate many other Multi-View clustering approach.

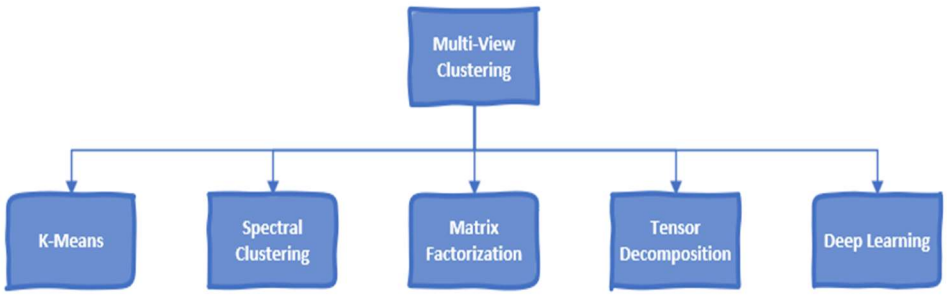

Figure 2: Multi-View clustering methods (Source- Self)

2. LITERATURE REVIEW AND RESEARCH GAPS

Numerous researchers did explore "Clustering - Unsupervised Machine Learning" however very few have specifically addressed "Multi-View Clustering". This field remains challenging for research since many facts are still unexplored hence, an extensive literature review has been carried out by examining the following papers:

S. Bickel and T. Scheffer [1] developed multi-view clustering algorithms for text data, focusing on both partitioning (k-Means, k-Medoids, EM) and hierarchical methods. They evaluated these algorithms using the WebKB and 20-newsgroup datasets. Their study found that Multi-View Expectation Maximization significantly outperforms single-view algorithms, while agglomerative multi-view clustering often yields equal or worse results. A noted limitation is that EM can result in empty clusters.

Chang X, Tao D, and Xu C [2] reviewed multi-view learning algorithms, classifying them into co-training, multiple kernels, and subspace learning. They explored integrating multiple views using consensus or complementary principles, employing various kernel methods. Using the Citeseer and WebKB datasets, they found that multi-view learning outperforms single-view methods. However, the strong assumption of conditional independence and lack of generalization error bounds are notable limitations.

Sun S. [3] reviewed theories and methodologies in multi-view machine learning, identifying open problems to encourage further research. They used Kernel CCA and sparse CCA for analysis, with paired datasets for CCA and labeled/unlabeled data for co-training. The study highlighted key theories and open research questions but had limited discussion on practical implementation challenges and specific real-world applications.

Zhao J, Xie X, Xin X, and Sun S. [4] reviewed theoretical foundations and recent developments in multi-view learning, categorizing methods into co-training, co-regularization, and margin-consistency styles. They used datasets like Handwritten Digit, Multi-PIE, and CUFSF for their analysis. Their findings include theoretical

advancements, and the introduction of PAC-Bayes bounds and Rademacher complexity for generalization error analysis. However, they noted limitations such as overfitting on small training sets and ignoring specific statistical properties of individual views.

Ali Mamdouh E, Yang S, and Xiaodong H. [5] proposed a content-based recommendation system using deep learning to enhance recommendation quality and scalability. They used a multi-view deep learning model for user and item features and a generative model for discovering common clusters across domains. Their datasets included Bing search logs, Bing News browsing history, Windows AppStore downloads, and Xbox view logs. They achieved up to 49% improvement for existing users and 115% for new users, with a Mean Reciprocal Rank (MRR) of 0.427. However, traditional algorithms like SVD and CCA struggled with large-scale data, while the deep learning approach scaled well.

Blaschko MB and Lampert CH [6] proposed correlational spectral clustering for image clustering with paired data, generalizing spectral clustering to multiple modalities. Using kernel canonical correlation analysis (KCCA) and optimizing through Rayleigh quotient and Schur complement decomposition, they evaluated their method on datasets like Israeli-Images, a multimedia image-text web database, and UIUC-ISD. Their approach outperformed traditional spectral clustering, showing empirical improvements. However, linear CCA did not improve results in some datasets, and noise levels can affect clustering performance.

Kailing K, Kriegel H-P, Pryakhin A, and Schubert M [7] introduced a superior clustering method for multi-represented data, effectively combining multiple representations. They used the union-method for sparse data and the intersection-method for large, unspecific clusters. Using protein data from the Swissprot database and text documents with TFIDF vectors, they demonstrated meaningful groupings. However, multi-instance objects were mainly classified, not clustered, and the lack of class labels hindered quality measurement.

Chaudhuri K, Kakade SM, Livescu K, and Sridharan K [8] investigated how two ‘views’ enhance clustering tractability using Canonical Correlation Analysis (CCA).

They applied CCA to the VidTIMIT database for audio-visual speaker clustering, finding improved separation conditions. However, their approach is limited by data distribution constraints and clustering algorithm assumptions.

Yin Q, Shu W, He R, and Wang L [9] proposed a novel multi-view clustering framework based on sparse subspace representation, developing an iterative algorithm for efficient solutions and theoretical analysis. They explored different co-regularization forms and integrated link priors for link-constrained clustering. Using datasets like UCI Handwritten Digit, 3-Sources Text, WebKB, Movies617, and Animal, they achieved state-of-the-art accuracy and normalized mutual information. No specific limitations were mentioned in the paper.

Jain AK's [10] work summarizes clustering methods, challenges, and emerging research directions. They address issues in data clustering using methods like K-means, DBSCAN, probabilistic models, subspace clustering, and hierarchical divisive clustering. Their research highlights challenges such as defining clusters, feature selection, and normalization. Clustering algorithms vary in objective functions, generative models, and heuristics, identifying high-density regions separated by low-density areas. Limitations include difficulties with high-dimensional data, the influence of data representation on performance, and the need for benchmark data and application integration.

Maldonado S, Carrizosa E, and Weber R. [11] proposed Kernel-Penalized K-means (KP-Kmeans) for feature selection while maintaining cluster structure. They iteratively adjusted kernel shape to reduce feature space dimensionality, using datasets like Iris, Wine, and various cancer diagnosis datasets. Their method outperformed alternatives in feature selection and clustering performance. However, few feature selection methods exist for unsupervised learning, and the proposed method shows limited stability for higher values of k.

Liang D, Zhou P, Shi L, Wang H, Fan M, Wang W, and Shen Y-D [12] developed a robust k-means algorithm using 2,1-norm in feature and kernel space, proposing the novel robust multiple kernel k-means (RMKKM) algorithm. They used datasets like YALE, JAFFE, ORL, and COIL20, demonstrating RMKKM's superiority on

benchmark datasets. They also investigated the Nyström method and random Fourier features for efficiency. However, outliers can cause large errors in standard k-means, and kernel method performance heavily depends on kernel choice.

Liu X, Li M, Wang L, Dou Y, Yin J, and Zhu E [13] proposed a joint imputation and clustering algorithm for incomplete kernels, achieving better clustering performance by integrating both processes. They conducted extensive experiments on datasets like Flower17, Flower102, Caltech102, and CCV, demonstrating superior performance, especially with high missing ratios. However, existing MKC algorithms struggle with incomplete base kernels, and two-stage approaches that impute before clustering can affect performance.

Wang S, Gittens A, and Mahoney MW [14] analyzed the Nyström method for scalable kernel k-means clustering, proposing a provable approximate solution. They used the Nyström method for kernel approximation, linear k-means, and spectral clustering, evaluating on MNIST and PenDigits datasets. Their findings show kernel k-means with Nyström outperforms spectral clustering, providing a 1+ relative-error trace norm guarantee for low-rank approximation. However, traditional Nyström lacks a relative error bound and may not satisfy the PCP property for clustering.

Zhang R and Rudnicky AI [15] proposed an efficient large-scale clustering scheme for the kernel K-Means algorithm, overcoming its limitations for large corpora. They used a disk-based strategy and three common kernel functions, evaluating on the MNIST dataset. Their scheme saved over 90% of running time and proved more efficient than traditional K-Means. However, kernel K-Means is limited by computation and storage costs, and the kernel matrix size may exceed memory capacity.

Dhillon IS, Guan Y, and Kulis B [16] unified weighted kernel k-means and spectral clustering under a single framework, showing that spectral clustering is a special case of kernel k-means. They used datasets like human fibroblast gene expression, yeast gene expression, and Pendigits. Their findings reveal that weighted kernel k-means can minimize normalized cuts without eigenvectors, validating its

effectiveness in gene clustering and handwriting recognition. However, k-means struggles with non-linearly separable clusters, and spectral clustering has computational overheads due to eigenvector calculations.

Tzortzis G and Likas A [17] estimated view weights in multi-view clustering, controlling sparsity for optimal performance. They used kernel-based weighted multi-view clustering, including MVKKM and MVSpec, and analyzed the effect of the p exponent. Using datasets like handwritten digits and Fourier coefficients, they found that weighted kernel combinations enhance clustering accuracy and reflect data properties effectively. However, noise can negatively impact accuracy, and the quality of views' information may be overlooked.

Cai X, Nie F, and Huang H [18] proposed a robust multi-view K-means clustering method for large-scale heterogeneous data integration. They used a structured sparsity 2,1-norm for robustness to outliers and an optimization algorithm for non-smooth objectives. Evaluating on datasets like SensIT Vehicle, Caltech101-7, and SUN 397, their RMKMC method outperformed baselines, showing robustness, stability, and efficiency. However, graph-based methods are sensitive to kernel choice and parameter tuning, and heavy computation limits their scalability.

Zhao H and Yun F [19] proposed the Dual-regularized Multi-view Outlier Detection (DMOD) method to detect class and attribute outliers in multi-view data. Using datasets like Iris, Breast, Ionosphere, and Letter, they demonstrated DMOD's effectiveness and superior performance compared to state-of-the-art methods. However, the method's limitations include a non-smooth (l_{2,1})-norm and binary integer elements in the (G(i)) matrix.

Chen X, Xiaofei X, Huang JZ, and Ye Y [20] introduced TW-k-means for Multiview data clustering with variable weighting, comparing it to five other algorithms. TW-k-means computes weights for views and variables, enhancing clustering performance. Using datasets like Water Treatment Plant and Yeast Cell Cycle, they found TW-k-means outperformed other algorithms and identified important variables. However, traditional methods are unbalanced due to variable and view importance, and TW-k-means variances are unstable with small k values.

Yu-Meng X, Wang C-D, and Lai J-H [21] propose the WMCFS algorithm for multi-view clustering and feature selection, validated through experiments on real-world datasets. They use the MVKKM algorithm for weighted kernel combination, the Co-regspec method to minimize view differences, and the LLC-fs algorithm for single-view clustering with feature selection. They test on the Mfeat dataset (handwritten digits with six views), the Reuters dataset (multilingual documents in six topics), and the Corel dataset (2000 instances in five image-based classes). Their findings show that WMCFS enhances clustering with feature selection, significantly improving multi-view clustering methods. However, high-dimensional data requires feature selection for better results, and existing single-source algorithms are not directly applicable.

Bo J, Qiu F, and Wang L [22] propose a weighting strategy for multi-view clustering with an efficient optimization algorithm, supported by theoretical analysis and experimental evaluation. They use simultaneous weighting on views and features, embedding a gradient-based optimization algorithm into k-means for weight computation, and compare their method with five state-of-the-art algorithms on real-world datasets. They test on a synthetic dataset, the Amsterdam Library of Object Images (ALOI), the Multiple Feature (MF) dataset, and the 3-Sources text dataset from BBC, Reuters, and The Guardian. Their findings show that the proposed method excels in multi-view clustering, with an efficient gradient-based algorithm for clustering centroids and weight vectors, and theoretical proof of convergence for updating weight vectors. However, challenges include incompatible views and high-dimensional data, with no explicit limitations mentioned in the paper.

Xu C, Tao D, and Xu C [23] develop the Multi-View Self-Paced Learning (MSPL) algorithm to enhance clustering accuracy, designing a smoothed weighting scheme for example and view complexities. They compare hard and smoothed weighting schemes using the Handwritten Numerals and Animal with Attributes datasets. Their findings show that the MSPL algorithm outperforms other methods in clustering accuracy and purity, especially on the Handwritten Numerals dataset.

However, nonconvex algorithms are prone to bad local minima with outliers, and it's challenging to separate examples and views into 'easy' and 'complex' categories.

von Luxburg U [24] introduces spectral clustering algorithms, discussing their advantages and disadvantages. They derive these algorithms from different graph Laplacians, expressing commute distance using graph Laplacians and eigenvectors, and relate unnormalized spectral clustering to commute time embedding. Using a toy dataset with 200 Gaussian points, a fully connected graph, and a 10-nearest neighbor graph, they explore the relationship between Ncut and random walk transition probabilities. Their findings detail various graph Laplacians, their properties, and practical aspects of spectral clustering. However, unnormalized spectral clustering can fail to converge to meaningful solutions, and small entries in eigenvectors can cause issues.

Hagen L and Kahng AB [25] approximate the optimal ratio cut partition cost using the second smallest eigenvalue. They develop fast Lanczos-type methods for heuristic ratio cuts and demonstrate effective clustering methods based on second eigenvector computation. They propose a spectral ratio cut partitioning heuristic for circuit netlists, utilizing the Lanczos method for sparse symmetric eigenvalue problems and EIG1/EIG1-IG spectral algorithms for ratio cut partitioning. Tested on industry benchmark suites, including MCNC and Hughes benchmarks, their findings show that the second eigenvalue effectively approximates the optimal ratio cut partition cost, and fast Lanczos-type methods successfully compute heuristic ratio cuts. However, they do not mention iterative improvement methods for partitioning algorithms, and intuitive models may not remain accurate as circuit sizes grow.

Shi J and Malik J [26] propose a novel approach for perceptual grouping in vision by introducing the normalized cut criterion for image segmentation. They use this formulation for graph partitioning based on global impressions, exploring its relationship with average cut and average association algorithms. Efficient computation is achieved using a generalized eigenvalue problem. Tested on the OpenVC dataset, their method optimizes the normalized cut criterion for image

segmentation. However, minimal spanning tree methods lack global impressions, and normalized cut minimization is NP-complete for grid graphs.

Kumar A and Daumé H [27] develop a multi-view spectral clustering algorithm based on the co-training concept, comparing it with baselines on synthetic and real-world datasets. They use a co-training algorithm and a Canonical Correlation Analysis (CCA) approach for clustering, testing on synthetic data with three views from a Gaussian mixture model, BBC news articles, and BBCSPORTS data. Their findings show consistent improvements in clustering performance, with the co-trained multi-view spectral clustering outperforming baselines, especially on Reuters multilingual data and BBC datasets. However, the algorithm assumes the same underlying clustering across views and does not address scenarios with multiple natural clustering.

Xia T, Tao D, Mei T, and Zhang Y [28] propose Multiview Spectral Embedding (MSE) to explore complementary properties of multiple views, addressing spectral embedding challenges effectively. They evaluate MSE using a Corel image gallery subset and a modified 20-newsgroup dataset. Their findings show that MSE effectively explores complementary views for low-dimensional embedding and performs well on multimedia data with complementary features. However, MSE struggles with nearly independent views and requires (O(N^3)) time for large (N), posing time constraints.

Nie F, Li J, Li X, et al. [29] develop a parameter-free framework for multiview clustering and semi-supervised classification, automatically assigning optimal weights to graphs without parameters. They formulate the semi-supervised classification problem as convex for a global optimal solution, achieving comparable performance to state-of-the-art methods on real-world datasets. Using the AMGL framework, they test on datasets including handwritten numerals, CiteSeer, NUS-WIDE, and MSRC-v1. Their findings show that the AMGL framework achieves optimal weights for multiple graphs and is applicable for multiview clustering and semi-supervised classification, ensuring global optimal results with a convex objective. However, no explicit weight factors are defined, and

the choice of parameter (g) is crucial for optimal performance.

Chen C, Ng MK, and Zhang S [30] formulate an optimal subspace for clustering within and across graphs, comparing block spectral clustering methods with other algorithms and addressing challenges in handling multiple graphs of different sizes. They use cluster-based spectral clustering, a normalized cut approach, inexact preconditioners for nonsymmetric matrices, and Schwarz iterations for symmetric positive semi-definite problems, tested on datasets like TestK2-123 and TestK3-456. Their findings show that the proposed block spectral clustering method outperforms existing methods in accuracy, with block Laplacian matrices enhancing clustering accuracy and computational efficiency. However, costly optimization and tensor computation remain challenges, as does consistency control across different graphs.

Wang Y-X and Zhang Y-J [31] systematically summarize NMF principles, models, and algorithms over five years, constructing an integrated framework and discussing open issues, modifications, and extensions. They use Newton-type approaches, gradient methods, and reduced quadratic models, tested on the OpenVC dataset. Their findings show that NMF enhances data compression and interpretability but has open questions regarding statistical underpinning, uniqueness, and formulations. NMF requires global optimization techniques and proper initialization methods, with generalized NMF extending its applicability. However, local minimization affects convergence, and the initialization of (U) and (V) influences solution quality.

Kivinen J and Warmuth MK [32] compare Gradient Descent (GD) and EG algorithms for linear prediction, introducing new algorithms based on Winnow for linear regression. They use GD for weight updates and propose EG*, which uses gradient components in exponents. They analyze algorithms using a potential function (V(w)), squared Euclidean distance, and relative entropy, employing amortized analysis for worst-case loss bounds. Their findings show that EG* has smaller loss with few relevant input variables, while GD suffers increased loss with irrelevant variables. However, GD struggles with irrelevant inputs, and EG has logarithmic dependence on them.

Lee DD and Sebastian Seung H [33]investigate NMF for image generation and

semantic analysis, exploring non-negativity constraints for parts-based representations and applying NMF to encyclopedia articles for semantic feature discovery. They use NMF, PCA, and VQ methods on a facial image database and Grolier encyclopedia articles. Their findings show that NMF learns parts of faces and text features, providing localized, parts-based representations due to non-negativity constraints, unlike PCA and VQ. However, NMF is not suitable for complex object images and lacks syntactic part relationships.

Liu J, Wang C, Gao J, and Han J [34] develop an NMF-based multi-view clustering algorithm for common latent structures, proposing a normalization strategy to ensure meaningful and comparable clustering solutions. They use Joint NMF for multi-view clustering and MultiNMF for consensus matrix and cluster label computation, tested on synthetic datasets, 3-Sources Text dataset, Reuters Multilingual dataset, and UCI Handwritten Digit dataset. Their findings show the proposed algorithm's effectiveness through experiments on synthetic and real datasets. However, challenges include maintaining meaningful clustering solutions and higher computational complexity compared to standard NMF.

He X, Kan M-Y, Xie P, and Chen X [35] investigate user comments for Web 2.0 item clustering, developing the CoNMF framework for multi-view clustering with co-regularization. They use pair-wise and cluster-wise CoNMF paradigms and k-means clustering, tested on the Last.fm and Yelp datasets. Their findings show that CoNMF improves clustering with user comments, outperforming k-means and CoSC, and enhancing clustering performance with users' views. However, they note quality imbalance in data sources and limited studies on NMF for multi-view clustering.

Cai D, He X, Han J, and Huang TS [36] propose the Graph Regularized Nonnegative Matrix Factorization (GNMF) algorithm for data representation, respecting graph structure. They compare GNMF with state-of-the-art methods using gradient descent and multiplicative updating rules to ensure nonnegativity of (U) and (V). Tested on the COIL20 image library, CMU PIE face database, and NIST Topic Detection and Tracking corpus, their findings show GNMF's superior

performance in real-world problems, considering local invariance and providing convergence proof. However, model selection remains challenging due to unclear theoretical methods, and non-convex objective functions hinder global minima identification.

Zong L, Zhang X, Zhao L, Hong Y, and Zhao Q [37] propose MMNMF for multi-view clustering with local geometrical structure preservation. They investigate consensus manifold and coefficient matrix construction, comparing their algorithm with existing NMF-based methods. Using multi-manifold regularization and consensus construction, they test on datasets like Webkb, 3Sources, Handwritten, Cora, ALOI, and Politics-uk. Their findings show that MMNMF outperforms existing algorithms on benchmark datasets. However, NMF struggles to preserve local geometric structure, and GNMF underperforms in multi-view NMF variations.

Weihua O, Shujian Y, Li G, Jian L, Zhang K, and Xie G [38] explore discriminative and non-discriminative information in multi-view data, improving classification using cross entropy loss in non-negative matrix factorization (NMF). They use cross entropy loss to enhance classification performance, employing NMF-based approaches and a block coordinate descent algorithm, tested on datasets like Reuters, YaleFace, BBC, Cornell, Texas, Washington, and Wisconsin. Their findings show that the proposed MV-DLCSL algorithm outperforms state-of-the-art methods, with cross entropy loss yielding better results than mean square error. However, not all information is useful for classification, and focusing on discriminative data is crucial, as consistent and complementary info may not always improve performance.

Hidru D and Goldenberg A [39] develop EquiNMF for data integration with automated unsupervised parameter setting, proposing a graph-regularized multi-view NMF with equal view contribution. They test EquiNMF on the Butterfly and ORL Face datasets, finding it outperforms other NMF methods on multiview imaging data, showing consistency and robustness across various parameter settings. However, graph regularization may lead to over-constrained models, reducing expressiveness, and constraints can make it difficult to adjust (V) from initialization.

Kalayeh MM, Idrees H, and Shah M [40] develop NMF-KNN for image annotation with a weighted multi-view extension, addressing dataset imbalance and feature fusion issues. They use weighted multi-view NMF and a nearest-neighbor approach, tested on the Corel5K and ESP Game datasets. Their findings show that weighted NMF-KNN improves image annotation with query-specific generative models, enhancing tag transfer and addressing dataset imbalance. However, models can become obsolete with new data, and ad-hoc feature-fusion methods are commonly used.

2.1 RESEARCH GAPS

Table 1: Research Gaps

No Comparison With	No Exploration of	Limited Exploration of
Additional clustering algorithms [11]	Analysis of speedups and future directions [25]	Alternative spectral clustering relaxations [24]
Advanced large-scale multi-view methods [18]	Consistent abnormal outliers in individual views [19]	Feature selection methods in clustering [11]
Deep learning clustering methods [7]	Distance between view-specific matrices for classification [38]	Large-scale multi-view datasets [9]
Deep learning methods for performance evaluation [29]	Hierarchical models for parts learning in NMF [33]	NMF for multi-view clustering [35]
Diverse clustering algorithms [22]	Impairing outlier detection performance [19]	Real-world multi-represented data applications [7]
Other kernel clustering methods [14]	Large-scale circuit design partitioning [25]	Real-world multi-view datasets [37]
Other large-scale clustering schemes [15]	Large-scale datasets for scalability assessment [29]	Scalability, complexity, efficiency, robustness [6], [16], [18], [20], [21], [23], [29], [30], [34], [40]
Other linear prediction algorithms [32]	Nyström method and random fourier features[12]	Spectral clustering methods [27]
Other multi-view clustering algorithms [8]	Parts learning from complex image databases [33]	Varying learning rates [32]
Recent clustering algorithms [20]	View quality and insufficient sparsity control in view weights [17]	View construction [3]
State-of-the-art algorithms [23], [34]		

The research on Multi-View Clustering (MVC) reveals several key gaps, summarizes in (Table 1). There is ambiguity in regularization methods for multi-view Non-negative Matrix Factorization (NMF), inadequate theoretical foundations, and challenges with convergence guarantees and view-specific property handling. Scalability and efficiency remain major concerns, especially for large datasets and missing data. Existing methods also suffer from limited evaluation, lack of comparison with state-of-the-art and deep learning models, and insufficient real-world application validation. Furthermore, imputation is disconnected from clustering, and advanced techniques like indefinite matrices, hierarchical models, and outlier detection are underexplored, necessitating future research to address these issues.

3. PROBLEM STATEMENT

Multi-view clustering (MVC) has become an increasingly important area in machine learning, as it enables the integration of data from multiple perspectives to gain a more comprehensive understanding of complex datasets. Despite the progress made in MVC through methods such as k-means, spectral clustering, matrix factorization, tensor decomposition, and deep learning, several limitations and challenges persist. Existing techniques often struggle with scalability, handling missing data, and ensuring robust integration of multiple views. Furthermore, these methods frequently lack the computational efficiency needed for large-scale applications and are not easily adapted to distributed computing environments. To address these gaps, there is a need to develop novel MVC algorithms that can efficiently integrate multiple views and handle missing data, while ensuring scalability and computational efficiency. Such algorithms need to be evaluated against key performance metrics such as clustering accuracy, computation time, and their ability to generalize beyond experimental settings to practical, real-world applications.

4. DESCRIPTION OF BROAD AREA

Multi-view clustering (MVC) is a rapidly evolving area within machine learning and data mining, focused on clustering data that comes from multiple, distinct perspectives or "views." In many real-world scenarios, data is collected from different sources or representations (Figure 1), each providing a unique perspective on the same set of objects. For example, in image processing, an image can be represented by different features such as color histograms, textures, and shapes; in bioinformatics, gene expression data may be complemented by protein interaction data. The goal of MVC is to effectively integrate these diverse views to create a unified clustering model that reveals the underlying structure of the data more accurately than considering each view separately.

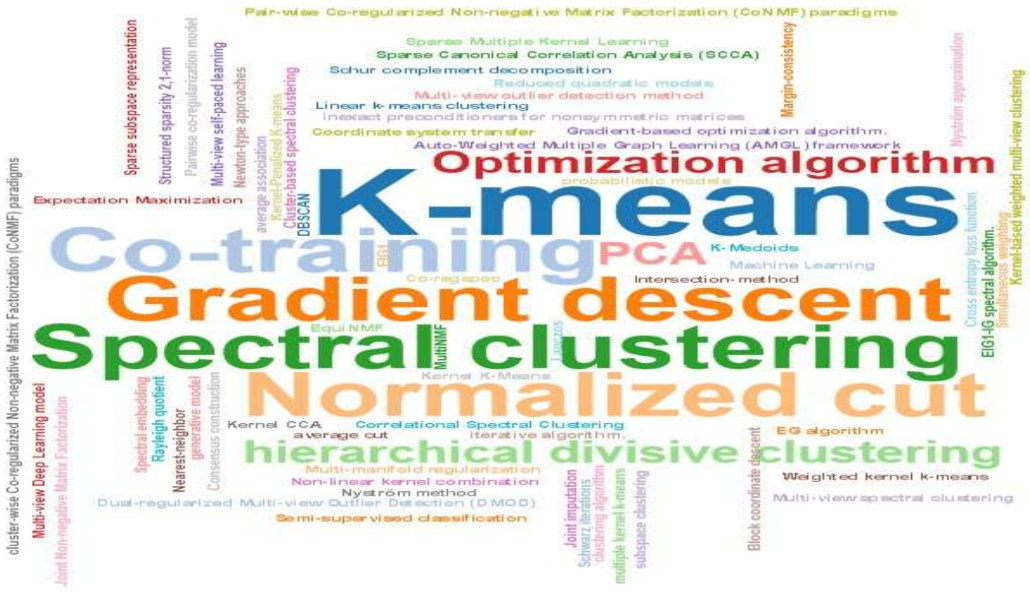

Figure 3: Word Cloud of Algorithm/Technique used (Source- Self)

To achieve this, a diverse range of algorithms and techniques are employed in MVC research, as illustrated by the Word Cloud (Figure 3). Core methods include K-means, Spectral Clustering, Gradient descent, and Kernel K-means, which are adapted for multi-view data through variants like Kernel-Penalized K-means and Weighted Kernel K-means. Non-negative Matrix Factorization (NMF) and its multi-view extensions, such as Co-regularized NMF (CoNMF), are used for data

decomposition and integration. Optimization techniques like gradient descent, block coordinate descent, and multiplicative updating rules are applied to ensure efficient convergence, while methods like Auto-Weighted Multiple Graph Learning (AMGL) and Canonical Correlation Analysis (CCA) help in fusing multiple views. These techniques collectively address challenges like data integration, scalability, and outlier detection.

Numerous techniques have been proposed to address the challenges of multi-view clustering, each offering distinct strengths and weaknesses. Building on the foundational work of Bickel and Scheffer [1], a variety of multi-view clustering methods have emerged. This research aims to review five prominent clustering methods and their multi-view adaptations: k-means, spectral clustering, matrix factorization, tensor decomposition, and deep learning.

K-means Clustering: One of the simplest and widely used clustering algorithms, k-means has been extended to multi-view settings. In multi-view k-means, the algorithm iterates between updating cluster assignments based on all views and updating the cluster centroids, ensuring that the final clustering reflects the information in all views.

Spectral Clustering: Spectral clustering techniques use the eigenvalues of similarity matrices to identify clusters. In multi-view spectral clustering, these similarity matrices are constructed for each view, and an aggregated matrix is used to find the optimal clustering. This method is powerful for capturing non-linear relationships between data points and is often effective in integrating multiple views.

Matrix Factorization and Tensor Decomposition: Matrix factorization techniques, such as non-negative matrix factorization (NMF), decompose data into a set of latent factors that can be used for clustering. These methods have been extended to multi-view settings by factorizing matrices representing each view and then jointly learning a shared low-dimensional representation of the data. Tensor decomposition is an extension of matrix factorization that can handle (Figure 4) more complex, multi-dimensional data structures, making it well-suited for MVC.

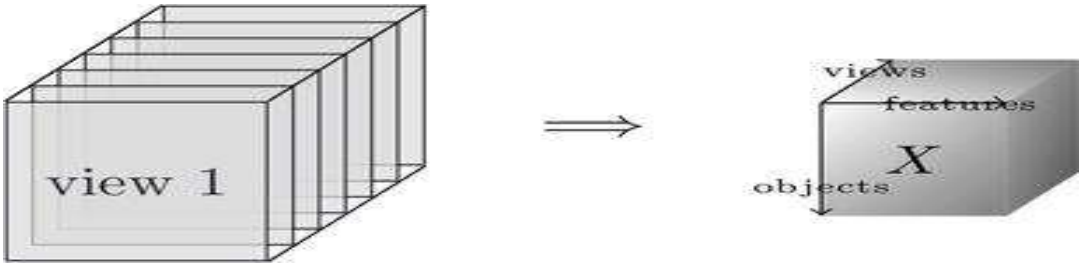

Figure 4: Matrix factorization techniques (Source- Self)

Probabilistic Models: Probabilistic models, including Bayesian and Gaussian mixture models, are commonly used in MVC to model the underlying data distributions for each view. These methods use probabilistic inference to combine information from multiple views and derive clustering assignments based on the joint probability of the data.

Deep Learning Approaches: Recently, deep learning techniques (Figure 5) have gained attention in MVC. Multi-view deep clustering methods often involve training neural networks to learn a common latent space or representation from multiple views. These methods leverage the powerful representation learning capabilities of neural networks to capture complex, non-linear relationships across views, improving clustering performance.

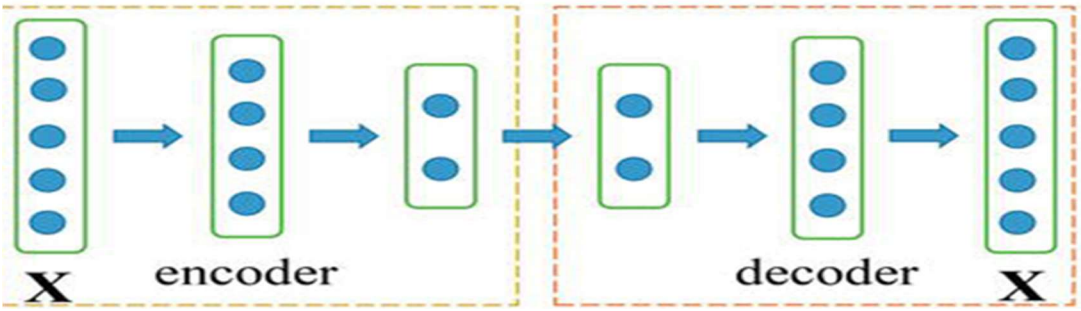

Figure 5: encoder, decoder (Source- Self)

5. RESEARCH OBJECTIVE

1. To critically review and analyze existing Multi-View Clustering (MVC) techniques.
2. To develop a novel MVC algorithm for effectively handling missing data in a dataset with multiple views.

3. To design a novel MVC algorithm that enhances the integration of multiple views through approaches such as distributed computing, regularization, and probabilistic modeling.
4. To comprehensively evaluate the performance of MVC algorithms by assessing key metrics such as clustering accuracy, computational efficiency, scalability, and robustness.

6. RESEARCH METHODOLOGY

The research will employ a systematic methodology to address the limitations and challenges identified in current multi-view clustering (MVC) techniques. Initially, a comprehensive literature review will be conducted to critically analyze existing methods such as k-means, spectral clustering, matrix factorization, tensor decomposition, and deep learning. This review will uncover gaps related to scalability, handling of missing data, and efficient integration of multiple views, setting the foundation for algorithm development. Following this, novel MVC algorithms will be developed, incorporating techniques like distributed computing, regularization, and probabilistic modeling to address the identified issues. These algorithms will be implemented using modern machine learning platforms such as Python, TensorFlow, and PyTorch, ensuring compatibility with large-scale datasets and real-world applications. The performance of the algorithms will be rigorously evaluated based on key metrics including clustering accuracy, computation time, scalability, and robustness in managing missing data. Additionally, a comparative analysis with existing methods will be carried out to validate the improvements. Iterative testing and refinement will be performed to ensure the algorithms are both theoretically sound and practically applicable. Figure 6 below illustrates the graphical representation of the research methodology.

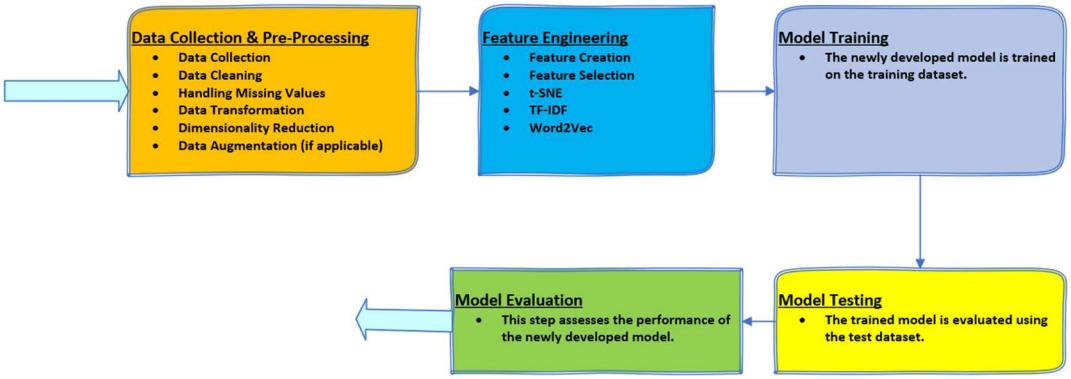

Figure 6: Research Methodology (Source: Self)

7. EXPECTED OUTCOME

The expected outcomes of this research are the development of novel, efficient multi-view clustering (MVC) algorithms that address current limitations in the field. Specifically, these algorithms will be capable of effectively integrating multiple data views and handling missing data using advanced methods like distributed computing, regularization, and probabilistic modeling. Additionally, these new algorithms are expected to outperform existing methods in terms of clustering accuracy, computation time, scalability, and robustness in dealing with incomplete data. This research will contribute significantly to the advancement of MVC, filling key gaps in the literature and offering practical, high-performance solutions for a range of applications in both academic and industrial contexts.

8. REFERENCE

[1] S. Bickel and T. Scheffer, "Multi-View Clustering," in *Fourth IEEE International Conference on Data Mining (ICDM'04)*, IEEE, 2004, pp. 19–26. doi: 10.1109/ICDM.2004.10095.

[2] C. Xu, D. Tao, and C. Xu, "A Survey on Multi-view Learning," Apr. 2013.

[3] S. Sun, "A survey of multi-view machine learning," *Neural Comput Appl*, vol. 23, no. 7–8, pp. 2031–2038, Dec. 2013, doi: 10.1007/s00521-013-1362-6.

[4] J. Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview: Recent progress and new challenges," *Information Fusion*, vol. 38, pp. 43–54, Nov. 2017, doi: 10.1016/j.inffus.2017.02.007.

[5] A. M. Elkahky, Y. Song, and X. He, "A Multi-View Deep Learning Approach for Cross

Domain User Modeling in Recommendation Systems," in *Proceedings of the 24th International Conference on World Wide Web*, Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, May 2015, pp. 278–288. doi: 10.1145/2736277.2741667.

[6] M. B. Blaschko and C. H. Lampert, "Correlational spectral clustering," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2008, pp. 1–8. doi: 10.1109/CVPR.2008.4587353.

[7] K. Kailing, H.-P. Kriegel, A. Pryakhin, and M. Schubert, "Clustering Multi-represented Objects with Noise," 2004, pp. 394–403. doi: 10.1007/978-3-540-24775-3_48.

[8] K. Chaudhuri, S. M. Kakade, K. Livescu, and K. Sridharan, "Multi-view clustering via canonical correlation analysis," in *Proceedings of the 26th Annual International Conference on Machine Learning*, New York, NY, USA: ACM, Jun. 2009, pp. 129–136. doi: 10.1145/1553374.1553391.

[9] Q. Yin, S. Wu, R. He, and L. Wang, "Multi-view clustering via pairwise sparse subspace representation," *Neurocomputing*, vol. 156, pp. 12–21, May 2015, doi: 10.1016/j.neucom.2015.01.017.

[10] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit Lett*, vol. 31, no. 8, pp. 651–666, Jun. 2010, doi: 10.1016/j.patrec.2009.09.011.

[11] S. Maldonado, E. Carrizosa, and R. Weber, "Kernel Penalized K-means: A feature selection method based on Kernel K-means," *Inf Sci (N Y)*, vol. 322, pp. 150–160, Nov. 2015, doi: 10.1016/j.ins.2015.06.008.

[12] P. Z. L. S. H. W. M. F. W. W. Y.-D. S. Liang Du, "Robust Multiple Kernel K-means Using `2,1-Norm", Accessed: Sep. 03, 2024. [Online]. Available: https://www.researchgate.net/profile/Peng-Zhou-39/publication/335079361_Robust_multiple_kernel_k-means_using_l21-norm/links/5d554d17299bf151bad41f63/Robust-multiple-kernel-k-means-using-l21-norm.pdf

[13] M. L. L. W. Y. D. J. Y. E. Z. Xinwang Liu, "Multiple kernel k-means with incomplete kernels", Accessed: Sep. 03, 2024. [Online]. Available: https://dl.acm.org/doi/abs/10.5555/3298483.3298565

[14] S. Wang, A. Gittens, and M. W. Mahoney, "Scalable Kernel K-Means Clustering with Nystrom Approximation: Relative-Error Bounds," Jun. 2017, Accessed: Sep. 03, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.1706.02803

[15] Rong Zhang and Alexander I. Rudnicky, "A Large Scale Clustering Scheme for Kernel K-Means," 2002, Accessed: Sep. 03, 2024. [Online]. Available: http://www.cs.cmu.edu/~rongz/icpr_2002.pdf

[16] I. S. Dhillon, Y. Guan, and B. Kulis, "Kernel k-means," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA: ACM, Aug. 2004, pp. 551–556. doi: 10.1145/1014052.1014118.

[17] G. Tzortzis and A. Likas, "Kernel-Based Weighted Multi-view Clustering," in *2012 IEEE 12th International Conference on Data Mining*, IEEE, Dec. 2012, pp. 675–684. doi: 10.1109/ICDM.2012.43.

[18] F. N. H. H. Xiao Cai, "Multi-view K-means clustering on big data," 2013, Accessed: Sep. 03, 2024. [Online]. Available: https://dl.acm.org/doi/10.5555/2540128.2540503

[19] Y. F. Handong Zhao, "Dual-regularized multi-view outlier detection," Jul. 2015, Accessed: Sep. 03, 2024. [Online]. Available: https://dl.acm.org/doi/10.5555/2832747.2832817

[20] Xiaojun Chen, Xiaofei Xu, J. Z. Huang, and Yunming Ye, "TW-k-means: Automated two-level variable weighting clustering algorithm for multiview data," *IEEE Trans Knowl Data Eng*, vol. 25, no. 4, pp. 932–944, Apr. 2013, doi: 10.1109/TKDE.2011.262.

[21] Y.-M. Xu, C.-D. Wang, and J.-H. Lai, "Weighted Multi-view Clustering with Feature

Selection," *Pattern Recognit*, vol. 53, pp. 25–35, May 2016, doi: 10.1016/j.patcog.2015.12.007.
[22] B. Jiang, F. Qiu, and L. Wang, "Multi-view clustering via simultaneous weighting on views and features," *Appl Soft Comput*, vol. 47, pp. 304–315, Oct. 2016, doi: 10.1016/j.asoc.2016.06.010.
[23] D. T. C. X. Chang Xu, "Multi-View Self-Paced Learning for Clustering," 2015, Accessed: Sep. 03, 2024. [Online]. Available: https://www.ijcai.org/Proceedings/15/Papers/558.pdf
[24] U. von Luxburg, "A tutorial on spectral clustering," *Stat Comput*, vol. 17, no. 4, pp. 395–416, Dec. 2007, doi: 10.1007/s11222-007-9033-z.
[25] K. A. Hagen L, "New spectral methods for ratio cut partitioning and clustering," 1992, Accessed: Sep. 03, 2024. [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=36278bf6919c6dced7d16dc0c02d725e1ed178f8
[26] Jianbo Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans Pattern Anal Mach Intell*, vol. 22, no. 8, pp. 888–905, 2000, doi: 10.1109/34.868688.
[27] H. D. Abhishek Kumar, "A Co-training Approach for Multi-view Spectral Clustering", Accessed: Sep. 04, 2024. [Online]. Available: https://dl.acm.org/doi/10.5555/3104482.3104532
[28] Tian Xia, Dacheng Tao, Tao Mei, and Yongdong Zhang, "Multiview Spectral Embedding," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 6, pp. 1438–1446, Dec. 2010, doi: 10.1109/TSMCB.2009.2039566.
[29] J. L. X. L. Feiping Nie, "Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification," 2016.
[30] C. Chen, M. K. Ng, and S. Zhang, "Block spectral clustering methods for multiple graphs," *Numer Linear Algebra Appl*, vol. 24, no. 1, Jan. 2017, doi: 10.1002/nla.2075.
[31] Y.-X. Wang and Y.-J. Zhang, "Nonnegative Matrix Factorization: A Comprehensive Review," *IEEE Trans Knowl Data Eng*, vol. 25, no. 6, pp. 1336–1353, Jun. 2013, doi: 10.1109/TKDE.2012.51.
[32] J. Kivinen and M. K. Warmuth, "Additive versus exponentiated gradient updates for linear prediction," in *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing - STOC '95*, New York, New York, USA: ACM Press, 1995, pp. 209–218. doi: 10.1145/225058.225121.
[33] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999, doi: 10.1038/44565.
[34] J. Liu, C. Wang, J. Gao, and J. Han, "Multi-View Clustering via Joint Nonnegative Matrix Factorization," in *Proceedings of the 2013 SIAM International Conference on Data Mining*, Philadelphia, PA: Society for Industrial and Applied Mathematics, May 2013, pp. 252–260. doi: 10.1137/1.9781611972832.28.
[35] X. He, M.-Y. Kan, P. Xie, and X. Chen, "Comment-based multi-view clustering of web 2.0 items," in *Proceedings of the 23rd international conference on World wide web*, New York, NY, USA: ACM, Apr. 2014, pp. 771–782. doi: 10.1145/2566486.2567975.
[36] Deng Cai, Xiaofei He, Jiawei Han, and T. S. Huang, "Graph Regularized Nonnegative Matrix Factorization for Data Representation," *IEEE Trans Pattern Anal Mach Intell*, vol. 33, no. 8, pp. 1548–1560, Aug. 2011, doi: 10.1109/TPAMI.2010.231.
[37] L. Zong, X. Zhang, L. Zhao, H. Yu, and Q. Zhao, "Multi-view clustering via multi-manifold regularized non-negative matrix factorization," *Neural Networks*, vol. 88, pp. 74–89, Apr. 2017, doi: 10.1016/j.neunet.2017.02.003.
[38] Y. W. R. L. X. L. Jian-wei Liu, "Multi-View Non-negative Matrix Factorization Discriminant Learning via Cross Entropy Loss," 2016.
[39] A. G. Daniel Hidru, "EquiNMF: Graph Regularized Multiview Nonnegative Matrix Factorization," 2014.

[40] M. M. Kalayeh, H. Idrees, and M. Shah, “NMF-KNN: Image Annotation Using Weighted Multi-view Non-negative Matrix Factorization,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2014, pp. 184–191. doi: 10.1109/CVPR.2014.31.

MANAV RACHNA
Vidyanatariksha

Dr. O.P. Bhalla Central Library
Manav Rachna International Institute of Research and Studies
(A deemed-to-be-university under section 3 of UGC Act 1956)

SIMILARITY VERIFICATION CERTIFICATE

Title: DESIGNING NOVEL MULTI-VIEW CLUSTERING ALGORITHMS FOR ENHANCED..... DATA INTEGRATION AND CLUSTERING PERFORMANCE.....

Researcher: PANKAJ KUMAR

Registration Number: 12301MRIIRS PHD0001

Supervisor: DR. RASHMI AGRAWAL

Department: COMPUTER APPLICATION

This is to certify that the above mentioned Short Synopsis/~~Thesis~~ was scanned for similarity index through iThenticate Software on 26 DEC 2024 with submission ID 3618374814414 and the similarity index was found to be ....8..... %. The complete report running into 25 Pages, 6,816 Word count with File Size 652.8 KB/MB is being forwarded to the Research Supervisor/DRC/DRB for review.

Date: 26 DEC 2024

Dr. Rajesh Kumar
University Librarian
Dr. O.P. Bhalla Central Library
Manav Rachna International
Institute of Research & Studies
Faridabad

The complete report of the above Short Synopsis/Thesis has been reviewed and the similarity index is below the accepted norms as per the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations 2018. The Short Synopsis/Thesis may be considered for further processing

26 DEC 2024

Research Scholar

Research Supervisor

Pankaj Kumar

Designing Novel Multi-view Clustering Algorithms for Enhanced Data Integration and Clustering Perfor

Manav Rachna International Institute of Research and Studies

Document Details

Submission ID
trn:oid:::3618:74814412

Submission Date
Dec 26, 2024, 4:19 PM GMT+5:30

Download Date
Dec 26, 2024, 4:27 PM GMT+5:30

File Name
Designing Novel Multi-view Clustering Algorithms for Enhanced Data Integration and Clusterin....docx

File Size
652.8 KB

25 Pages

6,816 Words

43,368 Characters

Dr. Rajesh Kumar
University Librarian
Dr. O.P. Bhalla Central Library
Manav Rachna International
Institute of Research & Studies
Faridabad

8% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Small Matches (less than 10 words)

Match Groups

- 20 Not Cited or Quoted 4%
  Matches with neither in-text citation nor quotation marks
- 20 Missing Quotations 4%
  Matches that are still very similar to source material
- 0 Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
6% Publications
0% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 20 Not Cited or Quoted 4%
  Matches with neither in-text citation nor quotation marks
- 20 Missing Quotations 4%
  Matches that are still very similar to source material
- 0 Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
6% Publications
0% Submitted works (Student Papers)

Rajesh Kumar
Dr. Rajesh Kumar
University Librarian
Dr. O.P. Bhalla Central Library
Manav Rachna International
Institute of Research & Studies
Faridabad

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
web.archive.org		3%
2	Internet	
bigdata.tsinghuajournals.com		1%
3	Internet	
arxiv.org		1%
4	Internet	
www.intechopen.com		0%
5	Internet	
www.geeksforgeeks.org		0%
6	Internet	
export.arxiv.org		0%
7	Internet	
www.arxiv-vanity.com		0%
8	Internet	
www.1staidsupplies.com		0%
9	Internet	
curin.chitkara.edu.in		0%
10	Publication	
Krzysztof Adamiak, Hyongsuk Kim, Krzysztof Ślot. "Accelerating projections to ker...		0%

11 Internet
addi.ehu.es 0%

12 Internet
www.novyimir.net 0%

13 Publication
Bo Jiang, Feiyue Qiu, Shipin Yang, Liping Wang. "Evolutionary multi-objective opti... 0%

14 Publication
Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, Wei-Ying Ma. "Collaborati... 0%

15 Publication
Xiaowei Gao, Xiaoke Ma, Wensheng Zhang, Jianbin Huang, He Li, Yanni Li, Jiangta... 0%

16 Internet
coek.info 0%

17 Internet
www.firstacademics.com 0%

18 Publication
"Neural Information Processing", Springer Nature, 2016 0%

19 Publication
Kun Song, Xiwen Yao, Feiping Nie, Xuelong Li, Mingliang Xu. "Weighted bilateral ... 0%

20 Internet
www.mdpi.com 0%

21 Internet
www.researchgate.net 0%

22 Internet
www.springerprofessional.de 0%